REMARKS

Responsive to the Office Action mailed 03 September 2009 and with an extension of time of three months, the present paper is timely filed on or before 03 March 2010.

By the present paper, claims 2 and 7 are amended and no claims are cancelled. Claims 2 - 9 and 11 are under examination.

Entry of the claim amendments and reconsideration of the Application are respectfully requested.

The Claim Amendments:

Claims 2 and 7 are amended to recite that the free energy of a combinatorial number of β-strands and α-helices is "performed in" polynomial time, consistent with the Office's construction of the phrase "as fast as polynomial time".

Applicant respectfully submits that the claim amendments do not introduce new matter into the Application.

Claim Rejections Under 35 U.S.C. § 112, paragraph second:

The rejections are based on the Office's interpretation or construction of various terms in Applicants' claims. Applicants acknowledge that the Office is required to construe a claim under examination as broadly as reasonably possible. However broad the construction, it must be conducted from the vantage point of the person of ordinary skill in the relevant art, and cannot be contrary to legal precedents. For example, a claim is definite if it reasonably informs the skilled artisan of the metes-and-bounds of the claims. The skilled artisan need only be able to determine with reasonable certainty when they were or were not infringing the claim.

Claims 2 -7 and 9 were rejected under 35 U.S.C. § 112,¶1, because the phrase "as rapidly as" allegedly renders the claims indefinite. Applicants respectfully traverse.

In Applicants' view, the Office implicitly, if not expressly, acknowledges that differences between "polynomial time" and "logarithmic time" are well-understood in the art. *See* Office Action at page 3, lines 1 - 6. The person skilled in the art, from whose vantage point the claim must be construed, understands that c(n+1)(n-1) describes a polynomial and that the calculation is thus performed as fast as well-understood polynomial time.

Without acquiescing to the Office's argument and solely to advance prosecution of the Application, Applicants have amended the claims to recite that the CLE model is performed in polynomial time, mooting the rejection.

Claims 2 and 7 were rejected because the constant "c" is not expressly defined, allegedly rendering the expression "c(n+1)(n-1) indefinite. Applicants respectfully traverse.

One of ordinary skill in the art understands algorithms and would possess general knowledge in treatises such as Robert Sedgewick, Algorithms in C, Addison-Wesley, Menlo Park, 1998 pp. 36-39. The cited text describes the difference in running times of algorithms without any explicit time in the expression. One of ordinary skill in the art would understand that the statement "as rapidly as" n^2 (i.e., even without Cn^2 or $O(n^2)$) is sufficient to mean that "the algorithm can calculate as rapidly as polynomial time and that polynomial time is of order n^2. The speed is polynomial time *for each repetition* of the algorithm. The real-time speed will still depend on implementation of the algorithm, which includes hardware limitations of the particular computer that is used, as well as on incidental programming matters (compilers, the length of the code, constant contributions from other modules, and the like). The constant "c" captures and accounts for these factors.

At page 15, lines 13 - 22, the instant specification discloses:

> Unlike the time it takes for the protein to fold, which is an exponential function of the global entropy (see example 5, Eqns (4) through (7)), the recursive search (without replacement) is done in polynomial time. If the secondary structures are all of the same size, then the time

> to align individual mers can be treated as a constant (A_t) times the recursive search (without replacement) of the secondary structure elements or $t_{search} \propto A_t B(N_{ss}+1)(N_{ss}-1) \propto N_{ss}^2$ where B is a constant and $\max\{A_t\} \propto \xi^3$ (depending on the procedure and assumptions used to evaluate A_t). This is a *substantial* gain over other methods. The underlying assumption in the CLE model is clearly that proteins don't wait a long time to find their fold. In the event that much longer folding times are thought to occur, a branch-and-bound approach is also an option.

The skilled artisan, knowledgeable in computer-implemented methods and models, would know that "c" involves software implementation issues (A_t) and hardware implementation issues (B). N_{ss} is the primary parameter (or "fixed power of the problem size" as acknowledged by the Office), where $N_{ss}(=n)$. Since ξ can vary in a calculation but the number of secondary structures units $N_{ss}(=n)$ does not, and since the discussion is directed to how these units come together, we thought to simplify the expression. For a given ξ and a given specification of the computer (hardware, motherboard, processor etc.), "c" is a constant, fixed by the architecture and operating software of the equipment chosen by one wishing to practice Applicants' inventive method.

The "trivial case" of c = 0 noted by the Office would require an infinitely fast computer. In this case, the particular algorithm used in the method would be irrelevant. The skilled artisan would know that a value of zero could not be assigned to "c". A claim is not indefinite simply because it encompasses a reasonably small number of inoperable embodiments that would be instantly recognized as such by the skilled artisan.

Applicants have merely chosen a term that is more concrete than the abstract "$O(n^2)$". *See* Office Action at page 7, line 5 (§ 103 rejection). The particular value of "c" would not be determinant of the issue of infringement *vel non* and thus its appearance in the claims does not render the claims indefinite. For at least these

reasons, Applicants respectfully submit that the rejection is improper and should be withdrawn.

Claim 9 was rejected under 35 U.S.C. § 112,¶2, because it is allegedly unclear what the variable ξ refers to. Applicants respectfully traverse.

Applicants respectfully submit that the skilled artisan knowledgeable in computer calculation and modeling of polymer structure, reading Applicants' specification, would understand ξ to denote persistence length, reflecting the "stiffness" of the polymer (polypeptide) chain. Persistence length, or the length of a segment of a hypothetical chain, has been understood in the art for over 60 years. *See* W. Kuhn and F. Grün, 101 Kolloid Z., 248 (1942). Equation (1) and the definition on page 9 line 7 are certainly within the framework of claim 9 and Equation (3) follows Equation (1) in claim 9. Since equation (3) follows equation (1) and both are in the same claim, it cannot possibly be misconstrued that ξ of Equation (1) is different from ξ in Equation (3). Therefore, the definition of ξ is not readily confused. The nature and meaning of ξ is described in the specification at, for example, pages 10 - 12 of the specification, in particular at page 12, lines 10 - 20.

As Applicants best understands the rejection, the Office bases the rejection, in large part, on the recitation that ξ can be replaced by a "dummy variable" "x". Applicants respectfully submit that "dummy variable" is well-understood in integral calculus and is not indefinite. As taught in Howard Anton, Calculus with Analytical Geometry, pp 331 - 332 (2nd ed., Wiley, NY, 1984) (formatting and emphasis supplied):

> Sometimes it is convenient to use a letter other than x for the variable of integration in a definite integral. For example, $\int_a^b f(t)dt$, $\int_a^b f(u)du$, $\int_a^b f(y)dy$. It is an important fact that: The value of a definite integral is unaffected if we change the letter used for the variable of integration, but do not change the limits of integration. As an illustration, let us compare the steps in the evaluation of $\int_1^3 x^2dx$, $\int_1^3 t^2dt$ and $\int_1^3 u^2du$. These integrals have the same limits of integration, but different variables of integration.
> $\int_1^3 x^2dx = x^3/3\Big]_{x=1}^{x=3} = (27/3)-(1/3) = 26/3$,

$$\int_1^3 t^2 dx = t^3/3\Big]_{t=1}^{t=3} = (27/3) - (1/3) = 26/3,$$

$$\int_1^3 u^2 dx = u^3/3\Big]_{u=1}^{u=3} = (27/3) - (1/3) = 26/3.$$ Because the letter used for the variable of integration has no effect on the final value of the definite integral, it is sometimes called a ***dummy variable***.

The Office Action states "[t]hus it is unclear what Applicants intend by redefining ξ, which in turn makes it unclear what ξ is." Applicants respectfully submit that they do nothing of the sort. The limits of integration in the integral are specified to be from $x = 1$ [amino acid] to x = ξ [amino acids] and (as expressed in the specification) that $1 \leq x \leq \xi$ and that the final result will depend only on the integration limits, which ultimately means that the function will depend on the two parameters that define the definite integral, namely x = 1 and x = ξ. Indeed, it is not particularly necessary to mention that x is a dummy variable to one of ordinary skill in the art". Applicants are unsure if the Office is suggesting that at least this basis for the rejection might be overcome by simply deleting reference to a "dummy variable".

Claim Rejections Under 35 U.S.C. § 103:

Claims 2 - 8 and 11 were rejected under 35 U.S.C. § 103(a) as allegedly obvious over Floudas et al., U.S. 6,832,162 (Floudas et al.) in view of Alm et al., 96 Proc. Natn'l Acad. Sci., 11305 - 310 (1999) (Alm et al.) in view of Dawson et. al., (Dawson et al. (2001)), and further in view of Turner et al., U.S. 5,424,963 (Turner et al.). Because the Office apparently misapprehends the disclosures of at least Floudas et al., Dawson et al. (2001), and Turner et al. and further because, in any event there was no motivation to combine Turner et al. with other applied art, Applicants respectfully traverse.

At column 10, lines 19 - 21, Floudas et al. discloses:

> These complications have led to the development of threading methods, an NP-complete (complexity related to performance of nondeterministic Turing machine in polynomial time) class of problems, in which the target sequence is threaded into the backbone of the template sequence while evaluating the sequence fitness.

As taught in Luger, G.F., Artificial Intelligence: Structures and Strategies for Complex Problem Solving, p156 (Addison-Wesley, 4th ed., 1997):

> The key conjecture in complexity theory is that there exists a class of inherently intractable problems. This class, referred to as NP (Non-deterministically Polynomial), consists of problems that may not be solved in less than exponential time without resorting to heuristics. Almost all search problems belong to this class.

The skilled artisan of the day understood that "NP-complete" means that the problem cannot in general be calculated in polynomial time. Floudas et al. are only defining the concept of "NP-complete". Applicants suspect that the unfamiliar nature of the word "non-deterministic" may have mislead the Office concerning the relevance *vel non* of the passage cited from Floudas et al. Nowhere does Floudas et al. disclose calculation in polynomial time. MD simulation, although using random movement through phase space, is also using a type of search which is clearly exponential, as Applicants have previous explained.

Neither Floudas et al. nor Alm et al. (discussed *infra*), nor Turner et al. (discussed *infra*) teach or suggest that the simulations therein disclosed can be addressed in polynomial time. They all suggest that the problem is "NP-complete", meaning that there is no way to ultimately reduce the time complexity from exponential growth.

Applicants turn next to Alm et al.

First Alm et al. do not apply a global correction to the entropy as Dawson et al (2001) do (discussed *infra*). Alm et al. apply a loop penalty entropy model that calculates entropy from the beginning position of the loop to the end position of the loop. A penalty applied to a short sequence forming a loop is defined as local entropy. Alm et al. do not consider every base pair interaction as Dawson et al (2001) do. In Alm et al., local entropy is a measure of entropy loss. This approach has a fatal flaw as demonstrated in Appendix C of the Declaration filed 15 February 2008.

Second, Alm et al. do not propose an algorithm that calculates in polynomial time, rather it is exponential time. It is not obvious from Alm et al. that a reduction in the degrees of freedom should be applied. Alm et al. do not mention persistence length or how this would influence their calculation. In fact, the lattice model and the cross-linking entropy model and the discredited loop penalty entropy of Alm et al. are incompatible with each other and cannot be reconciled without corrections, first recognized in Dawson and Kawai (2009). Moreover, unlike the loop penalty model, a lattice model is incompatible with a cross-linking entropy model. Finally, since Dawson et al. (2001) disclose that, given a structure, the cross-linking entropy can be calculated, and Dawson et al. (2001) was uninformed by the fallacious bias that these distorted prediction approaches yield (they were yet to be discovered), the model would have failed absent the understanding of Equation (2) and Equation (3) of claim 9 of the present invention.

Therefore, it was not obvious in what way Alm et al. could contribute to Dawson et al. (2001) (discussed *infra*) without additional information, unknown because of the distortions in information generated by the prevailing errant models of the day.

Applicants turn next to Dawson et al. and respectfully point out that the Office appears to overlook the significance of Equations (2) and (3) of instant claim 9 to Equation (9) of Dawson et al. (2001).

There is nothing attributable to the term $\sum_{ij} f_{ij}(\xi)$ in Dawson et al. (2001) because, as explained in Appendix 1 of the declaration under 37 CFR 1.132 filed 15 February 2008 where it is recited:

> Dawson et al did understand how the concept of a persistence length influences the global entropy. However, they did not understand that this must be accounted for at all levels of the calculation. They assumed, just like the authors of the sources where they obtained the structures to test their model, that they could manage to calculate the behavior of the RNA (and their claims about proteins) using a strictly monomer-by-monomer calculation scheme. They did not understand

> that they must fix their calculation strategy to account for the fact that the monomers (amino acids and nucleic acids) interact locally as a group in persistence-length related length-scales.

The effect of ignoring local interactions on RNA is shown in figures 1 and 2 of Appendix 1 to the Declaration filed 15 February 2008, These figures are quite different in appearance. The differences stem from that fact that the positive contributions $f_{ij}(\xi)$ discourage the formation of structures that do not conform to known constraints on the structure that follow from the magnitude of the persistence length. Thus, a beta sheet cannot be wrapped up in a circle shape, it must be basically straight, as one of ordinary skill in the art understood. Equation (1) of claims 8 and 9 and Equation (9) of Dawson et al (2001) contain no such restrictions whatsoever. Therefore, strange and inappropriate structures like Figure 2 of Appendix 2 of the Declaration filed 15 February 2008 are likely to occur when a prediction program is constructed only using Equation (1) of claims 8 and 9 and neglecting the last term in equation (2) of instant claim 9.

Dawson et al. (2001) did not envision and could not have envisioned that local structural corrections needed to be considered because calculations using the lattice model result in a contradiction. See Dawson, W. and Kawai, G. (2009), *Modeling the Chain Entropy of Biopolymers: Unifying two different Random Walk Models under One Framework*, J. Comput. Sci. Syst. Biol., 2:001-023 and Appendix 2 of the Declaration filed 15 February 2008. Calculations using the entropy model of Floudas et al (and other approaches like it) violate the second law of thermodynamics (see Appendix C of the Declaration of 15 February 2008).

Therefore, it was not obvious from Dawson et al (2001) that corrections to the local (global entropy) must be included. The authors of Dawson et al. (2001) did not have enough information, and the information they had was too flawed to allow one skilled in the art to deduce that there was a problem with the concept disclosed in Dawson et al. (2001).

Applicants reiterate that, whereas Dawson et al. (2001) disclose a method to calculate the entropy that is applicable to RNA and proteins, nowhere does the reference teach or suggest a way to predict RNA or protein structure. Rather, the authors assumed that using existing methods with their entropy correction would be sufficient. As demonstrated in Appendix 1 of the declaration filed 15 February 2008, this is not so.

Finally, whereas the approach disclosed in Dawson et al. (2001) correctly calculates entropy, Dawson et al. (2001) does not teach, suggest, or enable reducing the number of degrees of freedom. Dawson et al. (2001) only teach reduction in the number of degrees of freedom to describe and weight the magnitude of their entropy. The calculations still compute every base pair interaction. In Dawson et al. (2001), the final results of the entropic contribution from each computed base pair are weighted with the persistence length. Dawson et al. (2001) discloses that entropy must be computed for each base pair. Therefore, there is no relation between the approach of Turner et al. (discussed *infra*) and no reason for the authors of Dawson et al (2001) to have referred to the teachings of Turner et al.

Applicants turn finally to Turner et al. and respectfully submit that the Office misapprehends what Turner et al. would have taught or suggested to the person skilled in the art.

At column 7, beginning at line 44, Turner et al. recites:

> Multipole expansions permit a significant reduction of the time required to compute the $O(N^2)$ electrostatic terms. In the prior art, the evaluation of these terms typically takes 90% of the time required for the determination of the energy function. The multipole expansion replaces the static collection of N charges with M power series expansions where M<<N. By carrying out each expansion to a high enough order, the results are quantitatively similar to the exact calculations.

The term "multipole expansion" refers to the computation of electrostatic interactions in a given system where the number of electrostatic-charge interactions

between both positive and negative charges is such that the net charge difference is small compared to the number of charges (*See, e.g.*, J.D. Jackson, Classical Electrodynamics, chpt. 4 (2nd ed, Wiley, New York, 1975)) where multipole expansions are discussed in considerable detail).

Calculation of multipole interactions converges faster than computing the same interactions as point charges (electrostatic interactions). The computation of these electrostatic interactions between individual atoms is well known to one of ordinary skill in the art to require N^2 such computations. This is because one must calculate an expression $V = (1/2)\sum_j \left(\sum_{i \neq j} q_i q_j / (4\pi\varepsilon r_{ij}) \right)$ where $\sum_i 0$ means "to obtain the sum of all the items i in the list", V is the total electrostatic potential, q_i is the charge on atom i, q_j is the charge on atom j, r_{ij} is the distance between atom i and atom j and $4\pi\varepsilon$ expresses the dielectric constant of the system. a parameter that must be obtained by a measurement – or estimated – by one of ordinary skill in the art.

Thus, Turner et al. (col. 7, line 44-46) refers to the computation time of the electrostatic terms. It does not teach that "the evaluation of a combinatorial number of secondary structural element arrangements is determined in polynomial time defined by $o(n^2)$." One way the person skilled in the art might have performed this multipole expansion was to reduce the secondary structural units to a set of multipoles. This will reduce the time for computing the electrostatic terms. However, it has nothing to do with "evaluation of a combinatorial number of secondary structural element arrangements". It has to do with computing the electrostatic interactions at each time step of the molecular dynamics ("MD") simulation, where the MD simulation requires exponential time.

Secondly, Turner et al. teaches a method that helps significantly speed up and simplify the computation of molecular dynamics (MD) simulations for proteins. A brief summary of the purpose of the object of the invention is provided in the last two

paragraphs of column 3 (lines 58-68). In lines 58-65 of column 3, Turner et al. discloses:

> To study large systems these motions [sic.], as well as even large time scale motions, such as folding and unfolding transitions, a new class of dynamics algorithms is required that will reduce the degrees of freedom in a complex molecule system down to a manageable size, as well as scale up linearly for a complex system so that protein conformational changes can be studied at any desired level of detail.

Moreover, at column 2, lines 11-12 Turner et al. discloses, "[in] constrained dynamics approaches, constraints can be used to reduce the high frequency motion". The purpose is to reduce the number of degrees of freedom and to use MD simulation in the context of these reduced degrees of freedom.

The MD simulations involve computation of all the atoms in the form of bonding and non-bonding interactions. For short peptides (40 amino acids (aa)), standard MD simulations typically require an integration and sampling time of approximately 100 ns with a time increment of 2 fs (femtoseconds, 10^{-15}). Such would amount to about 10^8 time steps. Longer sequences and particular cases of even some shorter peptides can take a very long time to fold correctly. For example, there are instances of some proteins that take milliseconds (10^{12} time steps) or even seconds (10^{15} time steps) to completely fold. Since MD simulations are intended to accurately model this folding process, they are expected to also require the respective number of time steps to correctly do so. Because protein folding itself is not a controlled process and the folding times reflect a rate of folding for the majority of the population of that protein, there is no way to know at the beginning of an MD simulation how long this folding process will actually take because the simulation depends on searching random numbers to search and find the relevant conformations of the protein.

It can be further estimated that the number of time steps necessary will be on the order of q^N, where q is a constant and N is the number of amino acids. To be specific,

on page 169 in Figure 1 of Bryngelson et al (previously cited by the Office), Bryngelson et al. show a rough conceptual picture of the energy landscape for protein folding. Therefore, there are some natural effects in nature of the aa interactions that help the protein to fold faster than q^N, but the MD simulations depend on the use of random numbers to generate the conformation angles, and typical MD simulations involve the computation of these angles for atoms (i.e., not already dimensionally-reduced amino acids). As a result, the angle searches are not specified and require the cooperative effect of more than one such angle in a single amino acid, and, accordingly, the skilled artisan of the day would have understood that folding could take longer than q^N in some instances. Nevertheless, the skilled artisan of the day, given no other criteria, might have reasonably assumed that a simulation would take on the order of q^N time steps to complete.

Turner et al. discloses changing the value of N by reduction in the number of degrees of freedom (DOF) (*See* Office Action at page 7, line 8). Turner et al. therefore is using dimensional reduction from N atoms to some unspecified smaller number of interactions, which number of interactions is an adjustable parameter.

Applicants do not claim that they were the first to conceive of employing of reducing the number of degrees of freedom. The concept is employed in the instant inventive method, but not a limitation in the claims. Rather, Applicants' inventive method solves the problem of the predicting protein folding in polynomial time that is proportional to the number of secondary structure units. Applicants respectfully submit that the Office has conflated $O(N^2)$ in Turner et al. (the computation of the electrostatics of the atoms) with $o(n^2)$ (evaluation of a combinatorial number of secondary structural element arrangements). Nothing of that sort is taught or suggested in Turner et al.; that does not disclose even a computational time speed up time ($q^{\tilde{N}}$ where $q^{\tilde{N}} >> Cn^2$).

Accordingly, Applicants respectfully submit that the disclosure of Turner et al. is

not applicable to the task of "evaluating a combinatorial number of an arrangement of secondary structural elements". Turner et al. is directed to computing the electrostatic interactions at each time step of the MD simulation and gaining a significant speed up of q^N. The Office has not provided a reasoned statement why Turner et al. teaches "evaluation of a combinatorial number of secondary structural element arrangements" in polynomial time. For at least these reasons, Applicants respectfully submit that the rejection is improper and should be withdrawn.

Summary of Applicants' Traversal of Rejections Under 35 U.S.C. § 103:

Applicants respectfully submit that Dawson et al. (2001) at most discloses a vague idea that the cross-linking entropy might, in principle, be applicable to the problem of predicting or modeling protein structure. No teaching or suggestion of calculation in polynomial time can be found in Turner et al., Alm et al., or Floudas et al. Although dimensional reduction or "reduction in the number of degrees of freedom" was a concept known at the time, it is not an express limitation of Applicants' claims.

Alm et al. use a lattice model in the method therein disclosed and, moreover, employ an errant model of entropy as shown in Appendix 3 of the Declaration filed 15 February 2008. The present inventors have shown that the lattice model is inappropriate. Dawson, W. and Kawai, G., *Modeling the Chain Entropy of Biopolymers: Unifying two different Random Walk Models under One Framework,* J. Comput Sci Syst Biol. 2:001-023; Appendix 2 of the Rule 132 Declaration filed 15 February 2008. The "entropy penalty model" taught by Alm et al. violates the second law of thermodynamics as shown in Appendix 3 of the Rule 132 Declaration filed 15 February 2008). Applicants respectfully submit that it is improper to apply a reference that teaches or suggests a model that violates a law of nature. In any event a method based on a lattice model of folding functions in exponential time, q^N would not have made Applicants' inventive method for predicting folding functions in polynomial time obvious.

Further, both Floudas et al. and Turner et al. disclose MD simulations. Therefore,

neither approach teaches or suggests a polynomial time algorithm and they do not have a direct application to the computation of the global entropy in proteins or RNA. Entropy in MD simulations can only be inferred after doing a long MD simulation and then applying statistical sampling analysis of all the conformations that the protein moved through during that simulation. The skilled artisan of the day would have expected that this would require far more than the mere folding time to compute in order to obtain a statistically significant sample set of this data. At the time the present invention was made, there was no motivation to seek a way to calculate entropic parameters from a full scale MD simulation for at least the reason that, still today, it is in effect extremely costly. Hence, neither Turner et al. nor Floudas et al., alone or in combination with Alm et al. and Dawson et al. (2001) would have suggested to the skilled artisan calculating the thermodynamics or evaluating the entropy. Applicants respectfully submit that there was nothing obvious about applying a global entropy calculation to an MD simulation because the two are not even compatible.

Finally, Dawson et al. (2001) propose the global entropy. However, the reference fails to recognize the bias in the data sets therein employed. The reference teaches or suggest only Equation (1) and half of Equation (2) of instant claim 9. No teaching or suggestion of the second half of Equation (2) and the entirety of Equation (3) can be found in the art before 2006. Because the authors of Dawson et al. (2001) could not have seen the bias in their data set at that time, Dawson et al. (2001) could not have possibly taught or suggested the correction applied in the second half of Equation (2) of instant claim 9. Similarly Dawson et al. (20010 could not have taught or suggested local corrections to the global entropy (the local entropy). Both of these corrections can only be found in Applicants' disclosure.